Quarterly REPORT

(for July-September 1997)

Contract No. NAS596060

Enhanced Land cover and Land Cover Change products from MODIS
Algorithm Development and Post Launch Studies

by

John R. Townshend
Department of Geography
University of Maryland College Park
MD 20742

Submitted December 11th 1997

1. At-launch Land Cover Product.

a. Task Objectives:

i) The principal objective of this task is to supply a
validated at-launch land cover product based on the AVHRR at
a resolution of 1 km.

ii) A subsidiary objective is to gain agreement on the
classes to be used in the product.

b. Task Progress:

i) We successfully generated a 1 km land cover product to be
included with EDC's DISCover product as the MODIS at-launch
product.

Much of the coding was finished to prepare a Version 2 1 km
at launch land cover product. This product will conform to
MOD12 file specifications as much as possible and will
contain both the UMD and EDC land cover products. It will be
tiled and gridded in the integerized sinusoidal projection
and possibly the Goode's Homolosine projection as well.

UMCP convened a meeting in September to determine validation
procedures for the 1 km at-launch product. A report of this
meeting is provided in Appendix 1 to this report.

Preliminary results of the 1 km product generation were
presented at the 23rd Annual Meeting of the Remote Sensing
Society in Reading UK.

The possibility of using angular information to improve land cover classification was reported in a paper by Shunlin Liang and John Townshend at IGARRS 97 in Singapore in August.

ii) No activities carried out.

c. Anticipated Activities During the Next Quarter:

i) Our 1km map will be completed in December and open to within department reviews. A final version should then be available by the end of the year, possibly in the first or second week of January.

Finish coding for output of the 1 km product and deliver to SDST for use as simulated data.

Take part in a 1 km land cover validation exercise at the EROS Data Center Sioux Falls South Dakota in the first half of December.

ii) No activities anticipated.

2.   Land cover change indicator product.

a) Task objectives

i) Generation of test data sets

ii) Production and testing of the at-launch change detection algorithm.

iii) Production and testing of post-launch change detection algorithm

b) Task progress

i) The NALC Landsat MSS triplet images were collected and relevant change areas and change types were searched in order to prepare test data sets in addition to the previously prepared test data from Landsat TM imagery; TM image pairs of the BOREAS study sites were spatially rectified, and test data was created from same.

ii) Several change detection algorithms were coded into the MODIS Vegetation Cover Conversion product software which was delivered on time in August. The structures of the Look-Up Tables required for the change detection algorithms were determined. The algorithm code were successfully tested with the selected test data sets before delivered.

We began work on an experiment to examine seasonal and inter-annual variations in image texture measures for our post-launch product. These results will allow for refinement of look-up table thresholds values for labeling land cover

conversion.

iii) 1km and 8km AVHRR data were utilized in creating the
spectral trajectories of various theoretical change types.
These trajectories were employed in creating the structures
and contents of Look-Up Tables required for the change
detection algorithms. This provides a pilot procedure for
using the future real MODIS data to refine the Look-Up
Tables.

c) Anticipated Activities during the Next 3 months

i) All test data previously prepared will be rechecked for
quality and stored with uniform format for testing the coded
change detection algorithms.

ii) The computer code of the MODIS VCC product will be
examined in SDST. A complete set of the Look-Up Tables
required by the change detection algorithms will be obtained
using the AVHRR data sets/

iii) An automatic procedure for the creation and evaluation
of the Look-Up Tables of the change detection algorithms will
be carried out for the refinement of the post-launch MODIS
product.

3. Continuous fields of land cover properties

a) Task objectives

Generation of continuous fields of land cover attributes

b) Task progress

Using the AVHRR Pathfinder data as a prototype, we have
developed a methodology to provide continuous fields of three
types of continuous variables: (i) % bare, % herbaceous and
% woody, ii) % deciduous and % evergreen, iii) % needle-leaf
and % broadleaf. The methodology for locations of endmembers
involves the use of linear regression based on the training
data developed at UMD. The method has been applied to AVHRR
1 km data and products for each of the data layers have been
generated. We are currently conducting comparisons with
other land cover products, such as the EDC 1 km
classification, to assess the accuracy of the continuous
fields. Discussion is underway regarding the desirability of
including these data planes in the MODIS at-launch land cover
product.

c) Anticipated Activities during the Next Quarter

We plan to assess the accuracy of the prototype AVHRR product
for continuous fields based on comparisons with other land
cover products. Based on the outcome of the comparison, we

will refine the methodology as necessary. In addition, we plan to develop code for deriving continuous fields from MODIS data as a post-launch product.

MODIS AT LAUNCH LAND COVER MEETING

16 SEPTEMBER 1997

UNIVERSITY OF MARYLAND, UNIVERSITY COLLEGE CONFERENCE CENTER

Participants
John Townshend, UMD
Rob Sohlberg, UMD
Cheng Huang, UMD
Doug Muchoney, BU
Charlene DiMiceli, UMD
George Riggs, NASA
Richard Stenger, SDST
Joe Scepan, UCSB
Brad Reed, EDC
Tom Loveland, EDC
Allan Strahler, BU
Ruth DeFries, UMD
David Roy, MODLAND
Xiwu Zhan, UMD
Matt Hansen, UMD
Rama Nemani, Montana
Honghui Liu, Chinese Academy of Sciences

INTRODUCTION - (TOWNSHEND)

Purpose of meeting: To ensure delivery of an appropriate At-launch land cover product
    Review: EDC/Discover product and other land cover products
    Review training and testing data for validation.
    Possible selection of data planes. for inclusion (may be delayed).
    "Bake-off" proposed at MODLAND may not be best approach
    Possibility for multiple products?
    Product selected should be focused on user needs

EDC/DISCOVER - (LOVELAND)

Over-view of approach for EDC 1 km. and results
    Version 1.2 of global product is now available
    Conducted partially through funding from USGS, NASA, other agencies
    Focus: satisfaction of IGBP needs
    Derived from 1 km. AVHRR Pathfinder Land Data Set
    Used NDVI values April 92 - March 93
    May still be some bad composites (JT, Reed)

Action: Need to clarify which whether all groups are using the same 1 km data set (Tom Loveland and Matt Hansen)

    Clouds are not labeled though location of main residual cloud is known
    Procedure: flexible database strategy
    Processed on a continent-by-continent basis
    Legend development was done by independent teams
    Only Discover layers will be formally validated
    All source data, DEMs, disaggregation regions and final product are on-line

Planned database components:
        Olson global ecosystems; SIB, SIB2, BATS,
        USGS/Anderson, IGBP, UNESCO
Non-Goode georeferencing now available
IGBP data set is done and available on a web site
Other data layers are done; still being moved to WWW site
Per class accuracy will be completed for IGBP product
Improved urban layer may be added
20% of land area has at least one month of cloud contamination
seasonal regions classified per Olson, then binned to other schemes (like SIB)
        Process works for 90% of the classes; others are revised by hand
        There will be periodic revisions (current is ver. 1.2

OTHER LAND COVER PRODUCTS

UMD 1 km. Land Cover Product -Hansen

Method and results for version 0.1
        Same period as EDC, but all bands
        Classes modified to conform with IGBP
        Landsat data used to define training areas
        metrics: maximum annual; minimum annual; mean annual; annual amplitude;
greenest, month value; mean of six greenest months
trees: free of assumptions regarding distribution; explicit linkages between
        Decision-tree approach is used.
        Sawtooth and clouds can contaminate classification in some cases
        Mosaics are problematic; as is semi-arid agriculture
        Loveland: Boreal problem may be problem with IGBP definitions

Action: Hansen & Loveland to compare notes on class boundaries
Action: Hansen to isolate "bad" swaths and convey information to EDC
        Western Europe is an issue in both classifications
Action: Roy to will be asked to provide guidance on European classifications

Towards version 1.0
        improve metrics (8 mos. & sawtooth problem)
        more training areas
        address issue of mosaics
        derive reduced data set for simplified tree

Continuous Fields - DeFries

        Move toward proportional cover within a pixel
        Seven data planes
        woodiness
                woody
                herbaceous
                bare
        leaf type
                needleleaf
                broadleaf
        leaf longevity
                evergreen

deciduous
Methodology
stratify globe into three regions
collapse metrics via linear discriminant analysis
determine end members using linear regression
deconvolve with linear mixture model
Issues
validation/accuracy assessment
use in models
leaf type estimate for low % woodiness

Action:: DeFries to check Eigen vectors to see which relative importance of metrics contributing to continuous fields
Action:: UMD/MODIS Team Member D. Hall to coordinate re products for snow and ice

Central America Terrestrial Ecology Program - Muchoney

Database for validation and classification activities
Multiple classification schemes examined
450 sites selected, representing range of ecosystems and physiognomy
database populated from many sources
Variables include field observation, climate, vegetation structure, soils
Plots established in Honduras, Guatemala, various reserve areas
Target is 16 AVHRR pixels for each plot
Smaller plots for conservancy users
Classification driven by NDVI, slope/aspect from DEM, FAO soils data
several classifiers use (neural network and decision trees)
25 land cover labels can be aggregated to IGBP labels
Classification system is based on potential vegetation; no anthropogenic classes
Next steps
re-analyze some of the training data
expand method to Jornada
examine scaling issues
The focus will be on more training data, rather than improving accuracy with current training areas

MODLAND FEEDBACK

Nemani: Categories fine except for need for broadleaf category.

Action:: need for broadleaf crops category.

It was concluded that there was a crucial need for agricultural land cover classification and that it was unlikely that a purely coarse resolution remote sensing approach could be successful. It would rely on use of fine resolution data and also ancillary data such as agricultural statistics.

Action:: JT to ask other MODLAND members for list of parameterizations they will make based on land cover information

Roy: Vermote says they need Montana categories + desert & water for ver. 2.1 of the reflectance product.

Riggs for Dorothy Hall: need land cover to refine snow/ice products within forest and thus some measure of forest density will be required but probably not until year 2. Version 2.0 snow/ice does not rely on land cover

The atmospheric people also need information on land cover.

Action Strahler to provide contact name for land cover needs

At-launch land cover product should be same form as post-launch product
- Gridded, fixed, non-overlapping, earth-located tiles
  - currently 1200x1200 km tiles
  - 338 tiles contain land
- EOS-HDF with meta-data satisfying ECS/MODLAND requirements
- QA requirements
  - calibration & characterizations
  - validation
  - pre-launch (sensitivity analysis)
- Implications
  `
  - Need to validate the at-launch product
  - Label the at-launch with ECS-defined QA
- Metadata can be defined for either a granule or for a layer
- All MODLAND products will carry two bits of mandatory QA
- Will also summarize QA % quality by tile/granule

Action: Accuracy may be best presented at the "collection level" accuracy data
This would apply to an entire land cover at-launch layer

Validation methodology
- IGBP Working Paper #13 details methodology
- 850 samples have been located
- Land surface divided into 13 regions
  - used to locate partners for each region (3 per region)
- Classes of participants
  - validation regional advisors -- provide names
  - cooperating laboratories

High resolution verification data
- primarily TM data; some Spot -- collected same time period as 1 km. PAL
- Interpretation keys are being developed for the workshop
  - provide general resources for the interpreters
- 288 TM scenes received from USGS
- Spot scenes identified; in process

Schedule
- originally November 1997
- protoflight now set for December 1st-5th (North America only)
- complete workshop will slip until March/April 1998 - too late for MODIS at launch delivery schedule.

Development of methods
- Strahler: co-registration is a big issue

header corner points are not sufficient
interpreters may have to do some co-registration themselves
Hansen commented that there may be variation of class boundaries between interpreters
Strahler said that interpreters will provide both an IGBP class and additional
description (such as what they would call the cover)

Townshend stated that there would be a need for a dry run before North America meeting
in December

Strahler stated that the product will be per class and per region standard errors

Strahler stated that the BU validation activity will not be ready in time for at-launch

OTHER VALIDATION ACTIVITIES

UMD Training/Validation Data Set -- DeFries

Data set obviously biased in favor of UMD
Classified 169 (mostly MSS) scenes
Scenes chosen
cloud free data available
ancillary data
within 1 degree between classification agreement area
one pixel buffer was removed from the edge
380,000 1 km. pixels in full data set

Issues
mislabeling from ancillary source
misregistration of 1 km. data

BU Training Sites -- Muchoney

Central America
Jornada
New England in development
GW National Forest

Remote sensing land cover/land surface parameter model
Biophysical parameters
horizontal and vertical structure
function, physiognomy, environmental variables
anthropogenic factors
Key points
classification free
formalized parameter/feature extraction
ecosystem based
standardized plot data
formalized parameter database
globally applicable
May add data for two-layer canopies

Summary of Available Validation Data - Townshend

Available validation data

1) DISCover validation - partial for North America: in fact strictly only applicable to DISCover
2) UMD training and validation
3) BU training and validation
4) Confidence sites e.g., Humid Tropics Pathfinder forest/non-forest.

The IGBP-DIS Confidence level process being led by Strahler will not be ready in time for MODIS at launch delivery

Conclusions:

Only partial validation possible for at-launch product.

It will not be possible to have an established single data set which can be used for validation of the different products.

PRELIMINARY DEFINITION OF AT LAUNCH PRODUCT

Selection of land cover data planes

The first issue is whether there should be one data plane or multiple ones. It is not clear that we can or need to decide now.

If there is to be only one, then general agreement that it should be DISCover.

If multiple data planes are to be included, which ones are to be selected?

a) Multiple EDC planes
b) UMD classes
c) UMD Fields continuous fields (whether these should be included in the at-launch product was unclear though this was in the original proposal from UMD).

There was some discussion of how serious would be the change in file specification if multiple planes were to be included. Does a change in file specification allowing multiple planes require re-coding for downstream processes?

Action: Need to find out what the implications will be if additional planes added to BU land cover file specification.

Assuming additional planes could be included then at this stage **it seems desirable to plan on having multiple data planes,** since different planes will probably satisfy different users to varying degrees.

Metadata for validation

a) Agreed to include minimum ECS/MODLAND per pixel QA and tile meta-data.

b) There will be optional addition of other QA and metadata which must be according to the existing BU file specification. There was some discussion of the desirability of including tile meta-data on the accuracy of the classification. Although this could be derived per tile using global accuracy figures taking account of the actual proportions of cover types in a given tile there was some concern that this would indicate a spurious geographic level of precision of accuracy figures.

c) Collection-level meta-data should be prepared at regional and/or global levels

There also needs to be collection level metadata that explains the reasons for the multiple layers and their relative merits for different purposes.

d) There should be internal peer review within the group of the meta-data provided with any submitted data planes.

FUTURE ACTIVITIES TO COMPLETE DELIVERY OF AT-LAUNCH PRODUCT

Inter-meeting activities

Action: Obtain parameterization information from MODLAND members (John Townshend).

Action: UMD, EDC and BU should attempt to assess which errors are most significant in terms of use of data through parameterization and consider possible modification of products to optimize their use by MODLAND.

Action: UMD and EDC will generate revised land cover products: "final" version of products to be available by Dec. 10th 1997

Action: UMD and EDC will generate validation information using available sources.

Action: Requirements for Landsat 7 TM data should be established, based on Muchoney collation of MODLAND validation sites. A meeting needs to be organized between appropriate MODIS and Landsat personnel. For Landsat 7: Goward/Williams and MODLAND Justice and Muchoney). Muchoney to initiate.

ACTION: UMD and EDC should discuss product differences.

2nd At Launch Land Cover Product meeting

> Jan 13th -14th January at UMD.
> Review available products and validation
> Assess implications of product limitations on MODLAND algorithms
> Decide on submission of products for at-launch.

Delivery schedule

> Final delivery end of February
> Strahler stated that at some point will need full data to test PGEs.
> **ACTION: Townshend check implications of redelivery for SDST**: if
no major ramifications, then submit DISCover. May need to wait for completion of tiling scheme
> **ACTION: DiMiceli will reproject DISCover to integerized sinusoidal**
> **AGREEMENT: meeting for consideration of products in College Park**
> **January 13th and 14th**

ANY OTHER BUSINESS

Special journal issue

> include change detection
> stick to global and regional scale
> Loveland: interested in including DISCover
> **ACTION: DeFries will circulate letter soliciting interest**

The meeting was reconvened at 9 a.m. in the Laboratory for Global Remote Sensing Studies on the 17th September for informal discussions on differences between the two 1 km products.